**ChenWu: Explore NIPS papers dataset**

Summary of the report:

This project aimed to test different dimensionality reduction techniques on the NIPS papers data set to visualize any patterns in the data structure. PCA and MDS were applied to the data set and the results were visually compared. ISOMAP, LLE, and LE were applied to the K nearest neighbor graph and common authors graph of the data set and the results were also visually compared.

Describe the strengths of the report:

Linear dimensionality reduction techniques (PCA, MDS) and nonlinear manifold learning techniques (ISOMAP, LLE, LE) were applied to the data set, producing some intuitive visual representations of the different techniques.

Describe the weaknesses of the report:

Report contains some grammatical errors which hinders the clarity and quality of the writing.

Evaluation on clarity and quality of writing: 3

Minor grammatical mistakes hinder the comprehension of the work (see some examples below). The flow of the report will benefit greatly if these are cross-checked. This written work may be better conveyed in a poster format. Current format is difficult to follow, especially in pages 3 and 4 ( all the figures are grouped together and the interpretation is only found afterwards). A suggestion is to group the relevant analysis/interpretation with the corresponding figures.

Line 5: “artificial intelligence developed…” word choice error

Line 27: “in the report” spelling error

Line 56: ”We use PCA and MDS for…” word choice ‘to realize’ seems awkward

Line 58: “according to year” word choice error

Line 59: “Apply PCA and MDS…” may be a more suitable word choice than ‘use’

Evaluation on technical quality: 3

The authors tested many techniques and attempted to provide a relevant interpretation of the figures. As the data set contains a lot of textual information (i.e. title, pdf\_name, paper\_text etc.), it is unclear if this information was utilized. If so, there is no mention of any NLP techniques to enable such analysis.

Overall rating: 3

Confidence on your assessment: 3